

Reconnaissance hybride statistique-structurale de sinogrammes par système d'inférence floue

Adrien Delaye – Sébastien Macé – Eric Anquetil

IRISA - INSA
Campus Universitaire de Beaulieu
35042, Rennes Cedex, France

{Adrien.Delaye, Sebastien.Mace, Eric.Anquetil}@irisa.fr

Résumé : Dans ce papier, nous proposons une méthode hybride statistique-structurale originale pour la reconnaissance en-ligne de caractères chinois. Les caractères sont modélisés par des règles d'inférence floue combinant des informations morphologiques et contextuelles formalisées de façon homogène. Nous avons défini un ensemble de primitives élémentaires correspondant aux tracés qui peuvent intervenir dans l'écriture de caractères dans un style non contraint. Ainsi, chaque tracé analysé peut être classifié en primitive en évitant toute segmentation. Les règles d'inférence sont construites en couplant une information a priori sur les primitives constituant les caractères et une modélisation automatique de leur positionnement relatif au sein des caractères. Le système d'inférence floue agrège ces règles pour prendre la décision de classification. La méthode proposée a été validée par des premières expérimentations atteignant un taux de reconnaissance de 97.5 %.

Mots-clés : Reconnaissance en-ligne de caractères chinois, modélisation hybride statistique-structurale, système d'inférence floue, évaluation du positionnement relatif.

1 Introduction

La reconnaissance en-ligne de caractères chinois manuscrits a fait l'objet de recherches intensives durant ces trente dernières années [LIU 04]. Contrairement aux caractères latins, les caractères chinois peuvent être composés de nombreux traits fondamentaux (jusqu'à trente), organisés selon une structure bidimensionnelle (cf. Fig. 1). Outre la complexité des caractères, certaines caractéristiques de la langue chinoise telles que le grand nombre de classes de caractères (plus de 6000 utilisées au quotidien) ou la grande variabilité dans l'ordre de tracé et les variations de formes entre les différents styles d'écriture en font un problème de reconnaissance de formes très difficile. Les enjeux sont pourtant particulièrement importants, car les méthodes de saisie au clavier sont très inadaptées à la langue chinoise, du fait du grand nombre de classes de caractères.

Tout comme pour la reconnaissance de caractères latins, les méthodes mises en œuvre pour la reconnaissance de caractères chinois peuvent être regroupées en deux grandes catégories : les approches statistiques et les approches structurales. Dans les méthodes statistiques, le caractère à reconnaître est décrit de façon globale par un vecteur de carac-

téristiques, qui est classifié par une méthode de classification statistique standard : réseaux de neurones, plus proches voisins, séparateurs à vastes marges, ou fonction discriminante quadratique [YAN 07]. Dans ce type de méthode, la difficulté majeure réside dans le choix de caractéristiques robustes et discriminantes [LIU 07]. À l'inverse, les méthodes structurales utilisent l'information sur la structure hiérarchique des caractères chinois pour les décomposer en sous-formes, les modélisant par des primitives symboliques et des relations entre ces primitives. Par exemple, la méthode présentée dans [ZHE 97] représente un caractère par un graphe relationnel attribué flou dont les sommets modélisent les primitives et les arcs modélisent leurs relations. Les primitives utilisées dans les travaux de Zheng et al. sont les traits fondamentaux du chinois (lignes polygonales), mais on trouve ailleurs des modèles structurels reposant sur des segments de droite ou encore des sous-structures composées de plusieurs traits. La segmentation d'une séquence de tracés manuscrits en primitives est un problème commun aux méthodes structurales. Ce problème est particulièrement délicat lorsque les caractères sont écrits dans un style d'écriture courante (voir les exemples de la figure 1).

Dans ce papier, nous présentons une étude qui a pour but de démontrer la faisabilité d'une approche hybride structurale-statistique pour la reconnaissance de caractères chinois manuscrits s'appuyant sur une modélisation simple basée sur la théorie des sous-ensembles flous [ZAD 78]. Ce cadre théorique est particulièrement bien adapté pour modéliser des tracés manuscrits, qui sont par nature imprécis. Notre système a pour objectif d'être omni-scripteur et de permettre la reconnaissance de caractères tracés dans un style d'écriture non contraint (style courant).

Nous nous intéressons plus spécifiquement au sous-problème de la reconnaissance des radicaux du chinois. Les radicaux constituent un sous-ensemble des caractères chinois. Ce sont des caractères relativement simples, qui sont à la base de la composition de tous les autres. Tout caractère chinois, si complexe soit-il, peut être décomposé en ses radicaux, et le nombre de radicaux dans un caractère ne dépasse pas 5. Cette étude constitue donc une étape vers un système plus général pour la reconnaissance des caractères chinois.

Dans l'approche que nous présentons, la modélisation des radicaux est basée sur un ensemble de primitives qui permet d'éviter toute segmentation, en modélisant l'ensemble des

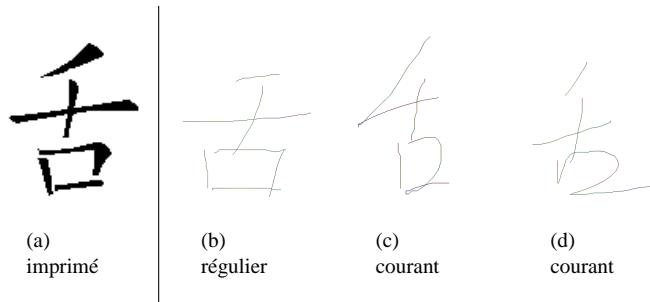


FIG. 1 – Formes imprimée (a) et variantes manuscrites (b,c,d) du caractère « SHE » dans différents styles d'écriture

types de tracés bruts pouvant intervenir lors de l'écriture de caractères chinois dans un style courant. Les radicaux sont décrits à partir de ces primitives au moyen de règles d'inférence floue modélisant des informations morphologiques (formes des primitives constituant le radical) et des informations liées au contexte spatial (positionnement relatif de ces primitives). La prise de décision du système de reconnaissance est faite en fusionnant ces deux types d'information au sein d'un Système d'inférence Floue. Cette démarche constitue une solution originale par sa simplicité : pas de segmentation, une description intuitive des caractères, mais qui aboutit à des premiers résultats prometteurs.

Dans la seconde section, nous introduisons les principes de l'approche et l'architecture générale du système de reconnaissance. La partie suivante présente le procédé d'apprentissage utilisé pour la modélisation des primitives et de leurs positions relatives. La méthode d'évaluation du positionnement relatif des primitives est décrite en détail dans la section 4. Ensuite, nous présentons l'exploitation de cette modélisation et le mécanisme de prise de décision en section 5. Enfin, une dernière section est consacrée à la présentation des premiers résultats expérimentaux.

2 Modélisation et architecture générale

Nous présentons ici les principes de la modélisation adoptée et l'architecture globale du système de reconnaissance.

2.1 Primitives de modélisation

Les caractères chinois sont théoriquement constitués de *traits fondamentaux*, lignes polygonales élémentaires, répertoriés au nombre de 20. Par exemple, le caractère de la figure 1(a) est constitué de 6 traits fondamentaux. Dans une version manuscrite du même caractère écrit dans un style d'écriture régulier, on retrouve bien distinctement les mêmes traits fondamentaux (cf. Fig. 1(b)). En revanche, lorsque l'écriture est plus relâchée (style d'écriture courant), des traits fondamentaux théoriquement séparés sont liés les uns aux autres, les lignes polygonales deviennent courbes, et on constate d'importantes variations de forme (cf. Fig. 1(c) et Fig. 1(d)).

Retrouver la décomposition canonique en traits fondamentaux à partir de ces tracés manuscrits courbes est un problème de segmentation très complexe. Plutôt que d'es-

p_1	ノ	p_9	フ	p_{17}	ㄣ	p_{25}	ㄣ
p_2	丨	p_{10}	ㄣ	p_{18}	ㄣ	p_{26}	ㄣ
p_3	ㄣ	p_{11}	ㄣ	p_{19}	ㄣ	p_{27}	ㄣ
p_4	ㄣ	p_{12}	ㄣ	p_{20}	ㄣ	p_{28}	ㄣ
p_5	ㄣ	p_{13}	ㄣ	p_{21}	ㄣ	p_{29}	ㄣ
p_6	ㄣ	p_{14}	ㄣ	p_{22}	ㄣ	p_{30}	ㄣ
p_7	ㄣ	p_{15}	ㄣ	p_{23}	ㄣ	p_{31}	ㄣ
p_8	一	p_{16}	ㄣ	p_{24}	ㄣ		

FIG. 2 – Dictionnaire de primitives

sayer de segmenter ces tracés courbes en traits fondamentaux, nous proposons de les considérer directement comme les primitives indivisibles de notre modélisation structurale. Ce choix s'appuie sur le constat de la stabilité de ces primitives : elles sont similaires d'un scripteur à l'autre, et tout caractère écrit dans un style courant est constitué des mêmes types de traits.

Dans leurs travaux, Kim et al. [KIM 96] s'appuient sur ce type de primitives pour décrire les caractères chinois selon une approche structurale. Une technique de classification non supervisée y est utilisée pour construire un dictionnaire des primitives rencontrées. Nous avons choisi au contraire d'établir ce dictionnaire de primitives à la main, afin de pouvoir attribuer un sens à chacune d'elles. En effet, ces primitives interviennent dans la description des radicaux par des règles de décomposition. Or l'écriture des règles nécessite la connaissance de la nature des types de primitives qui les constituent. À partir d'une base d'apprentissage représentative, nous avons donc construit un dictionnaire de 31 primitives présentées dans le tableau de la figure 2. Les primitives numérotées de 1 à 20 correspondent à des traits fondamentaux du chinois, tandis que les primitives 21 à 31 correspondent à des traits créés par liaisons et déformations de plusieurs traits fondamentaux.

À partir de ce dictionnaire, on peut décomposer un radical R_i en un ensemble de primitives p_{i_j} :

$$R_i \rightarrow p_{i_1} p_{i_2} \dots p_{i_n} \quad (1)$$

De cette façon, le radical « SHE » présenté à la figure 1(a) peut être décrit par :

$$R_{SHE} \rightarrow p_1 p_2 p_2 p_8 p_8 p_{10}$$

Deux autres décompositions décrivent les deux variantes 1(c) et 1(d) de « SHE » :

$$R_{SHE} \rightarrow p_2 p_2 p_{13} p_{24}$$

$$R_{SHE} \rightarrow p_1 p_2 p_8 p_{25}$$

2.2 Modélisation des radicaux par un système d'inférence floue

La modélisation formelle des radicaux repose sur le cadre théorique de la logique floue [BOU 03]. La description de

chaque radical R_i est formalisée à partir des primitives qui le composent par une règle d'inférence floue de la forme :

$$\underbrace{(F_{i_1} \wedge \dots \wedge F_{i_n})}_{\text{formes}} \wedge \underbrace{(C_{i_1} \wedge \dots \wedge C_{i_m})}_{\text{contexte}} \Rightarrow R_i \quad (2)$$

où F_{i_j} sont des conditions sur les formes des primitives composant R_i , et C_{i_j} sont des conditions sur le positionnement relatif de ces primitives. Une condition se traduit par un degré d'adéquation d'une variable avec un prototype flou au sens de la théorie des sous-ensembles flous [ZAD 78]. L'apprentissage des prototypes flous de modélisation des formes et du contexte fait l'objet des paragraphes 3.1 et 3.2. Les conditions de morphologie et de contexte sont combinées par conjonction floue (T-norme), notée \wedge , et l'activation de la conclusion de la règle R_i est déterminée par implication floue.

Pour une séquence de tracés manuscrits composant un radical à classifier, la prise de décision est faite par mise en concurrence des règles d'inférence dans un système d'inférence floue complètement homogène. Une sélection préalable des règles permet d'éliminer les règles non cohérentes, de sorte que les règles réellement mises en concurrence comportent toutes le même nombre de primitives, soit autant que le nombre de tracés manuscrits dans la séquence à identifier. Les scores des règles résultant de la conjonction du même nombre de prototypes flous, ils sont donc directement comparables. Cette homogénéité dans le système d'inférence est rendue possible par le principe de définition des primitives qui permet d'éviter toute étape de segmentation.

$$\left. \begin{array}{l} (F_{a_1} \wedge \dots \wedge F_{a_n}) \wedge (C_{a_1} \wedge \dots \wedge C_{a_m}) \Rightarrow R_a \\ \dots \\ (F_{k_1} \wedge \dots \wedge F_{k_n}) \wedge (C_{k_1} \wedge \dots \wedge C_{k_m}) \Rightarrow R_k \end{array} \right\} \text{agrégation}$$

On classe toutes les règles en comparant leurs scores d'activation par une méthode d'inférence *max-produit* pour prendre la décision de classification.

2.3 Architecture

Le schéma de la figure 3 illustre les différents modules de l'architecture de notre système qui interviennent dans le processus de reconnaissance d'un radical. Le tracé manuscrit de « SHE » de la figure 1(b), qui servira d'exemple pour illustrer plusieurs concepts tout au long de ce papier, est présenté en entrée au système sous la forme d'une séquence de tracés. Un système de reconnaissance de primitives attribue à chaque tracé un score vis-à-vis de chaque classe de primitives. Une sélection de règles du système d'inférence floue est faite à partir de ces scores, puis les degrés d'adéquation de formes et de positionnement sont fusionnées relativement à chaque règle d'inférence. Enfin, la comparaison des scores des conclusions des règles permet de prendre la décision de reconnaissance.

3 Apprentissage

Des règles de décomposition des radicaux en primitives sont écrites à la main par un expert, sous la forme (1). Si un même radical peut être décomposé selon plusieurs variantes, celles-ci sont décrites par autant de règles. C'est là la seule

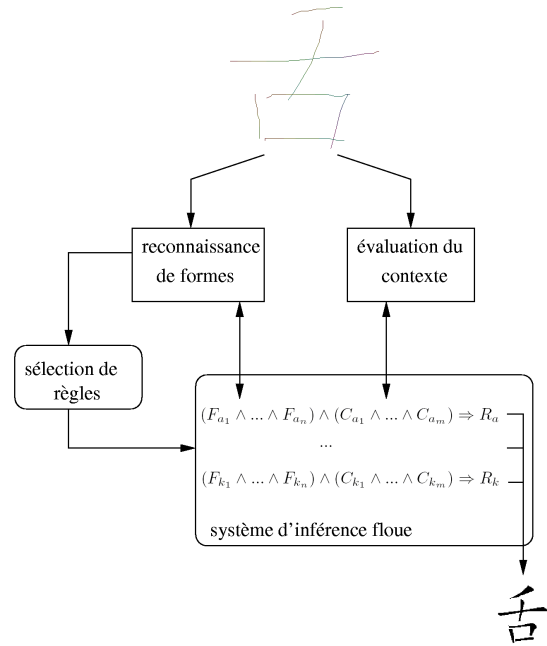


FIG. 3 – Vue d'ensemble du système de reconnaissance

intervention humaine nécessaire pour construire notre modèle : en effet, les règles d'inférence floue correspondantes, de la forme (2), sont générées automatiquement suite à l'apprentissage des prototypes flous modélisant les formes des primitives et de ceux modélisant leur positionnement relatif. Pour une décomposition $R_i \rightarrow p_{i_1} \dots p_{i_n}$ d'un radical R_i en n primitives, la règle d'inférence floue générée est :

$$\left. \begin{array}{ll} \text{formes} & X_1^f \in P_1^f \wedge \dots \wedge X_n^f \in P_n^f \\ & \wedge \\ \text{contexte} & X_2^c \in P_2^c \wedge \dots \wedge X_n^c \in P_n^c \end{array} \right\} \Rightarrow R$$

où les X_i^f et X_i^c représentent des variables liées respectivement à la forme et à la position du $i^{\text{ème}}$ tracé d'une séquence X à reconnaître. La prémisse de la règle est modélisée par n prototypes flous de formes (notés P_i^f) et par $n-1$ prototypes de positions relatives (notés P_i^c). La combinaison des prototypes de modélisation de positionnement permet d'évaluer à quel degré une séquence de tracés X respecte l'agencement attendu entre les primitives de ce radical R . Cet agencement est modélisé relativement à une primitive X_1 de référence pour les $n-1$ autres primitives de la règle.

Les prototypes flous P_i^f sont communs à toutes les règles puisqu'ils décrivent les classes de primitives introduites à la section 2.1. Au contraire, les prototypes flous P_i^c expriment un contexte de positionnement lié au radical, ils sont donc liés à la règle. Dans cette partie, nous décrivons la méthode d'apprentissage des deux types de prototypes flous, ainsi que les fonctions d'évaluation des degrés d'adéquation d'une variable avec ces prototypes.

3.1 Modélisation floue des classes de primitives

Les classes de primitives identifiées à la section 2.1 sont modélisées par des prototypes dont l'apprentissage est fait

de façon indépendante pour chaque classe : ainsi, la modélisation repose sur les propriétés intrinsèques caractéristiques des classes. Pour chacune, l'apprentissage se fait selon les deux étapes ci-dessous :

- chaque exemple d'apprentissage est représenté par un vecteur de 20 caractéristiques décrivant sa forme (courbure, dimensions relatives, nombre de zones descendantes, boucles...). Ces caractéristiques sont celles utilisées par RESIF, un système de reconnaissance d'écriture en-ligne pour des langues à alphabet latin [ANQ 97] ;
- les données sont regroupées dans l'espace de représentation par classification floue non supervisée (algorithme des C-moyennes floues) en prototypes définis chacun par une fonction à base radiale hyperellipsoïdale. L'ensemble des prototypes constitue une représentation qualitative des propriétés stables de la classe de primitives relativement aux caractéristiques choisies.

Pour une forme primitive de type inconnu, représentée par un vecteur X de caractéristiques extraites, on peut calculer un score d'activation relatif à chaque prototype flou P_r par :

$$\beta_r(X) = \frac{1}{1 + d_{Q_r}(X, \mu_r)} \quad (3)$$

où $d_{Q_r}(X, \mu_r)$ est la distance de Mahalanobis de X au centre du prototype μ_r , relativement à la matrice de covariance Q_r qui définit la forme du prototype. Le score de X pour une classe c est ensuite obtenu en fusionnant les scores des prototypes de la classe par un opérateur flou de type T-conorme (ici un \max). L'activation de la condition $F_j : X_j \in P_j^f$ s'évalue donc par :

$$F_j(X_j) = \max_{P_r \in P_j^f(j)} (\beta_r(X_j)) \quad (4)$$

où $P^f(j)$ dénote l'ensemble des prototypes flous participant à la modélisation de la classe j . Cette modélisation floue des primitives permet donc de décrire une forme inconnue par des scores exprimant des degrés d'adéquation vis-à-vis de chacune des classes modélisées.

3.2 Apprentissage des prototypes de positionnement flou

La méthode d'évaluation du contexte spatial que nous utilisons permet d'exprimer qualitativement la position d'une primitive par rapport à une autre primitive de référence (cf. section 4). La position d'une primitive X_j relativement à la primitive de référence X_1 est décrite par un vecteur $v^{X_1}(X_j)$. À partir de données d'apprentissage étiquetées, on apprend $n - 1$ prototypes pour chaque règle de n primitives. Un prototype $P_{i_j}^c$ modélise le positionnement relatif de la primitive i_j (par rapport à la référence i_1) dans la règle R_i .

Pour évaluer l'adéquation de la position relative de deux primitives par rapport à la position modélisée par un prototype flou, la fonction utilisée est la même que dans (3). L'activation de la condition de contexte $C_{i_j} : X_j \in P_{i_j}^c$ s'exprime donc par :

$$C_{i_j}(X_j) = \frac{1}{1 + d_Q(v^{X_1}(X_j), \mu)} \quad (5)$$

avec μ et Q les paramètres du prototype $P_{i_j}^c$.

4 Méthode d'évaluation du positionnement relatif flou des primitives

L'information sur le positionnement relatif entre les primitives d'un radical est cruciale pour sa reconnaissance. La plupart des méthodes structurales pour la reconnaissance de caractères chinois les décrivent par des propriétés catégorisées en types de relations telles que l'adjacence, l'intersection, le parallélisme, ... Pourtant, la position relative de deux traits manuscrits peut être trop ambiguë ou incertaine pour autoriser une prise de décision rigide à ce niveau. Plutôt que de forcer la catégorisation d'une relation de positionnement par ce type de procédé, nous proposons de nous appuyer là aussi sur une modélisation floue du positionnement relatif.

Nous exploitons l'approche introduite par Bloch dans [BLO 99] pour évaluer le positionnement relatif d'objets dans des images, adaptée aux tracés manuscrits en-ligne, à la manière présentée par Bouteruche et al. [BOU 06]. Cette méthode, qui repose sur la théorie de la morphologie mathématique, a l'avantage de mesurer des directions relatives de façon qualitative en utilisant le cadre de la logique floue. Elle permet d'évaluer avec quel degré d'adéquation une relation du type « être dans la direction D par rapport à R » est respectée par un objet relativement à une direction D et à objet R en référence.

Le principe d'évaluation de la position d'un objet analysé relativement à un objet de référence, tel qu'il est défini dans [BLO 99], peut être décomposé en deux étapes : la définition d'un *paysage flou* autour de la référence, puis la comparaison de l'objet analysé avec ce paysage flou.

4.1 Définition du paysage flou

Le paysage flou $\mu_\alpha(R)$, défini pour un élément de référence R relativement à une direction α donnée, est une fonction qui associe un degré d'appartenance flou à tout point de l'espace S , modélisant le degré avec lequel ce point vérifie la relation.

Notons P un point quelconque de S , et Q un point dans R . Soit $\beta(P, Q)$ l'angle formé par le vecteur \overrightarrow{QP} et la direction $\overrightarrow{w}_\alpha$, à valeurs dans $[0, \pi]$. $\beta(P, Q)$ est donné par l'équation :

$$\beta(P, Q) = \arccos\left(\frac{\overrightarrow{QP} \cdot \overrightarrow{w}_\alpha}{\|\overrightarrow{QP}\|}\right), \text{ et } \beta(P, P) = 0$$

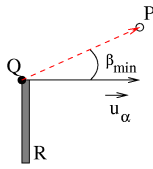
Pour tout point P de l'espace, on détermine le point Q de l'objet référence R tel que l'angle $\beta(P, Q)$ est minimal. On note $\beta_{\min}(P, R)$ cet angle. Ce mécanisme est illustré sur la figure 4.

Le paysage flou $\mu_\alpha(R)$ est alors défini par :

$$\mu_\alpha(R)(P) = f(\beta_{\min}(P, R))$$

où f est une fonction décroissante de $[0, \pi]$ dans $[0, 1]$. Nous choisissons pour f une fonction linéaire simple, ce qui aboutit à la définition suivante du paysage flou :

$$\mu_\alpha(R)(P) = \max\left(0, 1 - \frac{2\beta_{\min}(P, R)}{\pi}\right)$$

FIG. 4 – Définition de β_{min}

En un point P , le paysage flou μ_α représente donc l'adéquation de P avec la relation directionnelle définie par l'angle α par rapport à la référence R . Un point qui satisfait nettement la relation « être dans la direction \vec{u}_α par rapport à R » aura un haut degré d'appartenance au paysage. La figure 5 illustre des exemples de paysages flous : ils correspondent aux relations de direction « au dessus », « en dessous », « à gauche » et « à droite » par rapport à un trait manuscrit en référence, ici l'un des tracés du radical manuscrit de la figure 1(b).

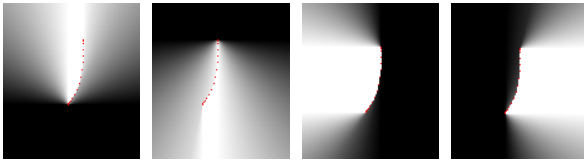


FIG. 5 – Paysages flous pour 4 relations de direction relatives à un objet de référence : « au dessus », « en dessous », « à gauche », « à droite » (les hauts niveaux de gris correspondent aux forts degrés d'appartenance)

4.2 Évaluation de l'appartenance d'un tracé à un paysage flou

Une fois le paysage flou défini, il reste à évaluer avec quel degré un objet analysé A respecte la relation considérée par rapport à l'objet de référence. Pour cela, nous utilisons la définition proposée dans [BLO 99] consistant à calculer la moyenne des degrés d'adéquation des points de A avec le paysage flou :

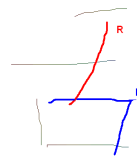
$$M_\alpha^R(A) = \frac{1}{|A|} \sum_{P \in A} \mu_\alpha(R)(P)$$

où $|A|$ est le cardinal de A .

4.3 Description du positionnement relatif des primitives

Le positionnement relatif d'une primitive d'un caractère chinois par rapport à une autre primitive en référence est évalué de cette façon suivant les 4 directions principales. Les relations considérées sont donc « en-dessous », « au-dessus », « à gauche » et « à droite ». La figure 6 présente un exemple de position relative mesurée pour un tracé (en bleu) par rapport à une référence (en rouge).

Aux quatre degrés flous qui décrivent qualitativement la position relative, nous ajoutons deux mesures des dimensions relatives de la primitive à positionner par rapport aux dimensions de la référence. Ce vecteur de positionnement $v^R(A)$



Direction	Degré
Haut	0.05
Bas	0.83
Gauche	0.21
Droite	0.86

FIG. 6 – Vecteur de positionnement relatif d'une primitive « P » par rapport à une référence « R »

est robuste face aux variations d'échelle et à l'étirement des caractères.

5 Exploitation

Dans cette section, nous présentons comment la modélisation définie précédemment est exploitée en phase de généralisation pour reconnaître des radicaux. Le processus de reconnaissance se fait en deux étapes : les règles d'inférence sont d'abord filtrées de façon à éliminer celles qui ne sont pas pertinentes. Puis les règles conservées sont mises en concurrence par le système d'inférence floue qui constitue le mécanisme de prise de décision.

5.1 Sélection des règles

La séquence de tracés $T = t_1 \dots t_n$ issue de la saisie manuscrite d'un radical est fournie en entrée au système de reconnaissance. Conformément à notre principe de modélisation, la segmentation de cette séquence n'est pas remise en cause, et donc les seules règles pertinentes sont celles qui comportent autant de primitives qu'il y a de traits dans T , soit n .

Les traits t_i sont comparés aux prototypes flous de modélisation des primitives, aboutissant au calcul des scores $F_c(t_i)$ de chaque trait t_i relativement à chaque classe de primitives c . Un seuil de rejet est fixé, qui détermine si la classe candidate c pour le trait t_i est pertinente ou non. Si pour une classe de primitive c , aucun trait t_i ne présente un score d'activation supérieur au seuil, toutes les règles constituées d'une primitive p_c sont éliminées. De cette façon, le nombre de règles à comparer est considérablement réduit.

5.2 Prise de décision par le Système d'Inférence Floue

Chacune des règles sélectionnées doit ensuite être évaluée vis-à-vis de la séquence de traits T . À ce stade, toutes les règles considérées comportent le même nombre de primitives. On peut donc attribuer des scores aux règles puis les comparer d'une manière homogène, formalisée par le SIF.

Pour chaque règle de la forme (2), n scores de formes, évalués selon (4), et $n - 1$ scores de position, évalués selon (5), sont fusionnés par un opérateur de T-norme floue pour établir un score global.

En pratique, un score est calculé pour toutes les permutations des traits d'entrée, pour faire face aux possibles variations dans l'ordre de tracé. Le score global d'une règle est donc finalement :

$$S(R_a) = \begin{matrix} F_{a_1}(u_1) & \otimes & \dots & \otimes & F_{a_n}(u_n) \\ \otimes & C_{a_2}(u_2) & \otimes & \dots & \otimes & C_{a_n}(u_n) \end{matrix}$$

où $u_i = t_{\sigma_{opt}^{-1}(i)}, \forall i \in \{1..n\}$, avec σ_{opt} la meilleure permutation sur $\{1..n\}$, c'est-à-dire maximisant le score $S(R_a)$. Les scores des différentes règles sont ensuite comparés et l'élément en conclusion de la meilleure règle est considéré comme le radical reconnu.

6 Expérimentations et résultats

Nos expérimentations ont été conduites sur un sous-ensemble d'une base de données constituée en Chine par le CASIA¹, qui regroupe 6763 classes de caractères écrits par 60 scripteurs différents, en style régulier et courant. Les exemples de caractères manuscrits présentés à la figure 1 sont extraits de cette base.

Le but de ces premières expérimentations étant avant tout de démontrer la validité de la démarche, nous nous sommes limités à la description de 35 radicaux, choisis parmi les plus courants. Ceux-ci sont constitués de 2 à 6 primitives, et décrits par 48 règles floues au total. L'apprentissage des prototypes flous de modélisation des primitives a été conduit sur un échantillon d'apprentissage de 1000 exemples par classe, écrits par 40 scripteurs. Parallèlement, les positions relatives des primitives au sein des radicaux ont été apprises avec 75 exemples pour chaque règle, écrits par ces mêmes 40 scripteurs.

Le test du système global a été conduit sur 1200 exemples de radicaux (25 exemples par règle), écrits par 20 scripteurs distincts de ceux ayant participé à l'apprentissage (cadre omni-scripteur), conduisant à un taux global de reconnaissance de 97.5%. Par la méthode de sélection des règles, le nombre moyen de règles considérées dans le processus d'inférence floue est abaissé à 12.3 règles pour une séquence d'entrée.

7 Conclusion et perspectives

Nous avons présenté un modèle hybride statistique-structurel pour la reconnaissance de radicaux du chinois. Une information a priori sur la structure des radicaux est injectée par l'écriture de règles de décomposition en primitives. Des règles d'inférence floue sont ensuite générées en incorporant des informations relatives aux formes des primitives et à leur positionnement relatif, suite à un apprentissage automatique. Ces informations morphologiques et contextuelles, décrites sous un formalisme homogène, sont fusionnées lors de la prise de décision par un Système d'Inférence Floue avec une méthode d'inférence max-produit. Grâce à la définition des primitives de modélisation, le système résultant est adapté à la reconnaissance de radicaux écrits dans un style courant, et aucune étape de segmentation n'est requise. Le système est également omni-scripteur, et insensible à l'ordre de tracé des traits. Cette approche a été validée par des premières expérimentations conduisant à un taux de 97.5% de bonne reconnaissance sur un sous-ensemble des radicaux du chinois.

Dans de futurs travaux, nous aurons pour objectif d'étendre cette méthode à la reconnaissance des caractères chinois au sens large, en s'appuyant sur la décomposition canonique des caractères en radicaux. Ce passage à l'échelle nécessitera une optimisation préalable des performances du système de

reconnaissance des radicaux. En particulier, il sera intéressant de définir un critère général pour identifier d'emblée la primitive de référence en phase de test, afin de diminuer la combinatoire engendrée par la recherche de la correspondance entre les traits d'entrées et les conditions floues des règles.

8 Remerciements

Les auteurs tiennent à remercier Cheng-Lin Liu, Professeur au CASIA, pour avoir mis à leur disposition les bases de données qui ont permis de mener à bien ces expérimentations. Merci aussi à Guy Lorette, Professeur à l'Université de Rennes I, pour ses précieux conseils. Ces travaux sont partiellement financés par la Région Bretagne.

Références

- [ANQ 97] ANQUETIL E., Modélisation et reconnaissance par la logique floue : application à la lecture automatique en-ligne de l'écriture manuscrite omni-scripteur, PhD thesis, Université de Rennes I, 1997.
- [BLO 99] BLOCH I., Fuzzy relative position between objects in image processing : a morphological approach, *IEEE Transactions on PAMI*, vol. 21, n° 7, 1999, pp. 657–664.
- [BOU 03] BOUCHON-MEUNIER B., MARSALA C., Eds., *Logique Floue, Principes, Aide à la Décision*, Hermès-Lavoisier, 2003.
- [BOU 06] BOUTERUCHE F., MACE S., ANQUETIL E., Fuzzy Relative Positioning for On-Line Handwritten Stroke Analysis, *Proceedings of the 10th International Workshop on Frontier in Handwriting Recognition (IWFHR'06)*, 2006.
- [KIM 96] KIM H., JUNG J., KIM S., On-line Chinese character recognition using ART-based stroke classification, *Pattern Recognition Letters*, vol. 17, n° 12, 1996, pp. 1311–1322.
- [LIU 04] LIU C.-L., JAEGER S., NAKAGAWA M., Online Recognition of Chinese Characters : the State-of-the-Art, *IEEE Transactions on PAMI*, vol. 26, n° 2, 2004, pp. 198–213.
- [LIU 07] LIU C., Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition, *IEEE Transactions on PAMI*, vol. 29, n° 8, 2007, pp. 1465–1469.
- [YAN 07] YANG D., JIN L., Kernel Modified Quadratic Discriminant Function for On-line Handwritten Chinese Characters Recognition, *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR'07)*, vol. 1, 2007, pp. 38–42.
- [ZAD 78] ZADEH L., Fuzzy sets as a basis for a theory of possibility, *International Journal on Fuzzy Sets and Systems*, vol. 1, n° 1, 1978, pp. 3–28.
- [ZHE 97] ZHENG J., DING X., WU Y., Recognizing On-line Handwritten Chinese Character via FARG Matching, *Proceedings of the 4th International Conference on Document Analysis and Recognition (ICDAR'97)*, 1997, pp. 621–624.

¹Chinese Academy of Science, Institute of Automation